# North Dakota County Fertility, Oil Prices, and Oil Output

David T. Flynn

Department of Economics & Finance
College of Business & Public Administration
University of North Dakota

This version: March 2019
Preliminary-Cite only with author's permission

**Abstract**

Hydraulic fracturing created an oil boom in North Dakota. The scale of the oil boom, and the additional economic effects created, altered and changed many aspects of life in North Dakota. I add to this literature by examining the changes in oil activity and their impact on the fertility rate in North Dakota. I undertake an examination at the state level first, and then refine the estimation to look at the county level fertility rate responses to the oil change. The state level analysis makes sense because the economic effects of the oil boom were not confined to oil producing counties alone. The county level analysis captures the fact that oil activity alone was specific to a defined subregion of the state. An interesting aspect of this is that significant increases in fertility in North Dakota occurred against the backdrop of steep declines in the U.S. as a whole.

## 1 Introduction

The ascension of oil into the status of a major industry in North Dakota was chronicled in the national news for shielding the state from recession (Davey, 2008), for generating labor market growth in excess of the region's capacity to house workers (Sulzberger, 2011), and for transforming many aspects of life in North Dakota (Brown, 2013). Despite the amazing stories and the clear changes in daily life, oil is not a new thing in North Dakota. The state records oil extraction going back to 1951. In fact, before the increased production made possible by hydraulic fracturing there was another, less pronounced, boom in production in the 1980s. It might be more appropriate to say that the *re–discovery* of oil in North Dakota created the massive shift in economic activity. When oil prices increased after 2000, and remained high, the industry in North Dakota responded with an enormous increase in production activity seen in Figure 1.

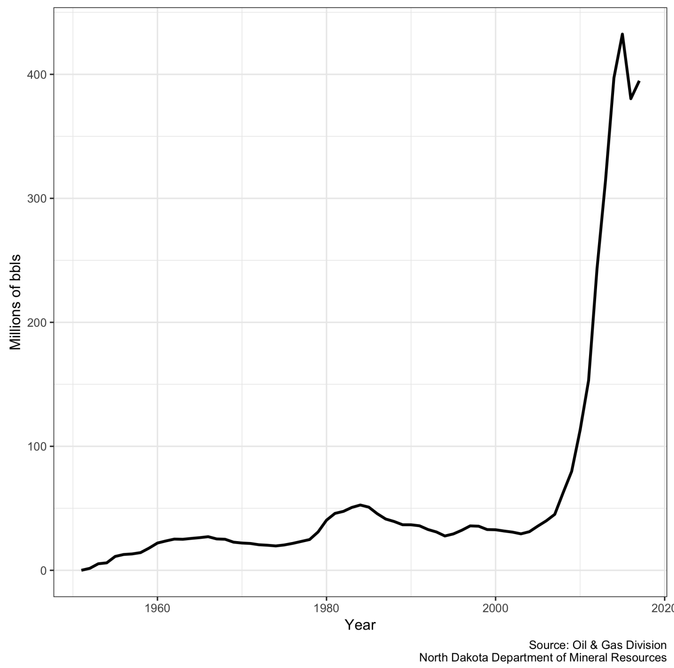


Figure 1: Monthly Oil Production in North Dakota

The dynamics of fertility rates and the implications of changes in fertility behavior are a topic of importance to academic research and the popular press. The preoccupation with fertility rates is, and has been, a global phenomenon (Tavernise, 2013). The addition of a further intersection, that of fertility behavior with economic performance, only adds to the interest. Becker (1991), with a chapter titled *The Demand for Children*, defined the economist's approach toward fertility decisions for a decades. Implementation of the approach is nuanced however because of the various data limitations encountered over time and across areas. For example, Roy et al. (1999) examined this issue for India and noted that preferences change over time and region and that even the definition of key variables such as wealth or income could be difficult to define in a universal manner. When the change in economic circumstances is sudden and sharp, an economic boom or bust, it might also change fertility behavior in a similar sharp or disjoint fashion, rather than smooth or gradual transitions.

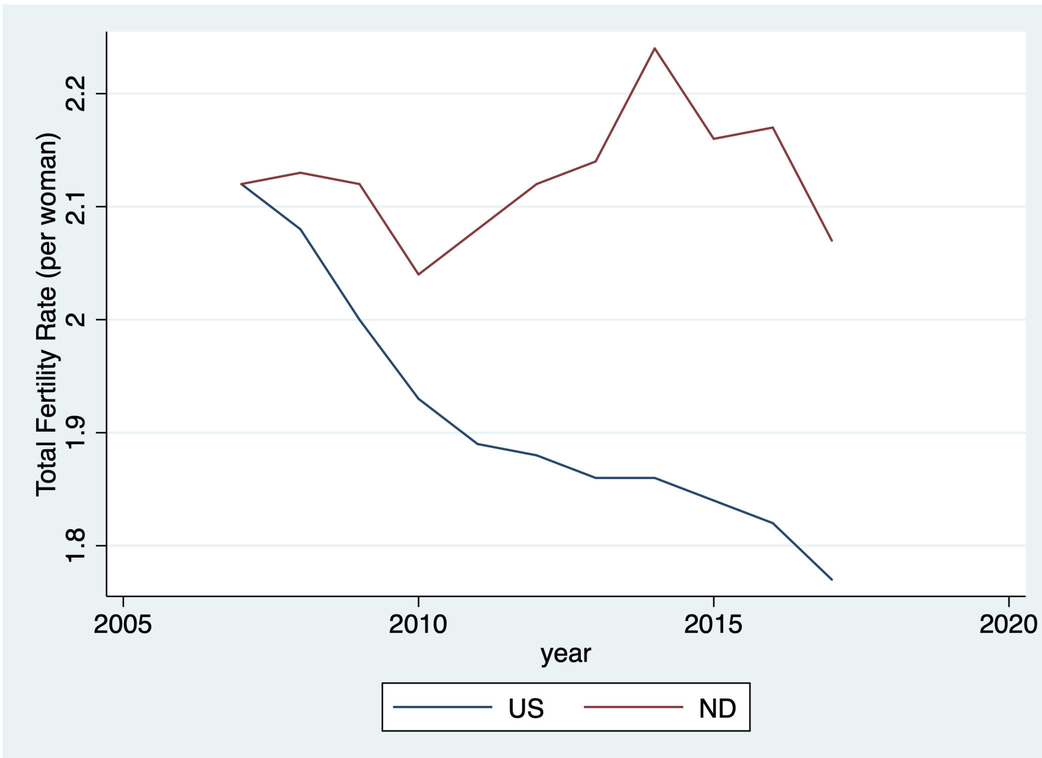


Figure 2: Total Fertility Rate, US &ND, 2007–2017

In the U.S. the recent focus for concern is declining fertility, recently to below replacement rates. North Dakota did not follow this pattern over the last ten years as seen in Figure 2.[1] In fact, the gap in fertility experience widened over the time, while North Dakota stayed, roughly, the same and the U.S. experienced a significant decline. I break out the fertility experience by age as well for 2007 and 2017 and compare the U.S. and North Dakota in Figures 3 and 4. The fertility rate of women between the ages of 20 and 30 is much higher in North Dakota than the U.S. in 2007. By 2017 both the US and ND rates shifted down,

[1]The data for the TFR come from various *National Vital Statistics Reports* such as Hamilton et al. (2015), Martin et al. (2016), and Martin et al. (2018).

though North Dakota fertility was much higher and extended its advantage even to age 35.

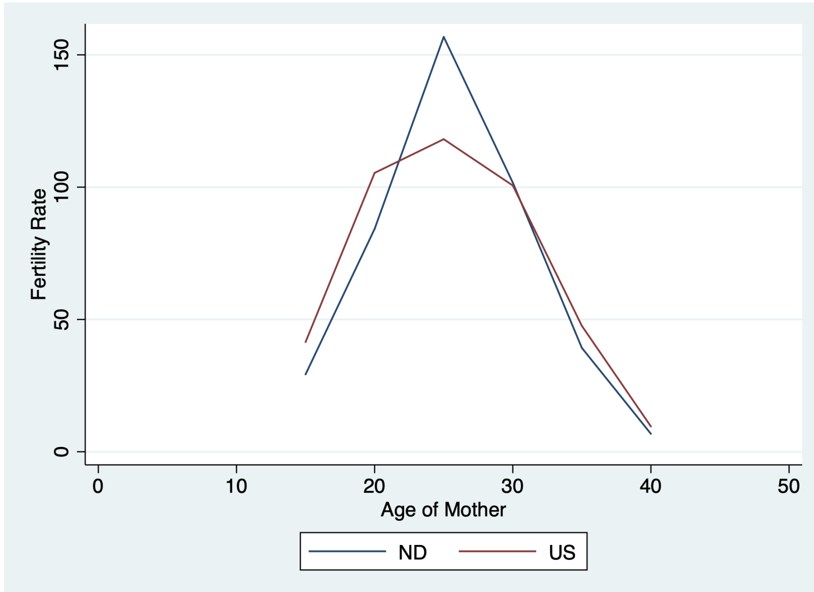


Figure 3: Fertility rate by age of mother, US & ND, 2007

The general approach of the paper highlights the literature on the connections between economic and social variables, such as fertility. In addition I look at literature examining either North Dakota or shale oil effects in general. After that I briefly discuss the data employed for the analysis in this paper and any manipulations made to get the data into usable forms. I then examine the fertility changes in North Dakota at the state level using oil variables as primary independent variable. After that I attempt to disaggregate the experience of North Dakota to the county level. While age–specific fertility rates are not feasible at this time for the county level, general fertility rates for counties can be calculated. In the state and county sections I discuss the results of the analysis, and then in the last section I conclude and discuss necessary modifications or extensions.

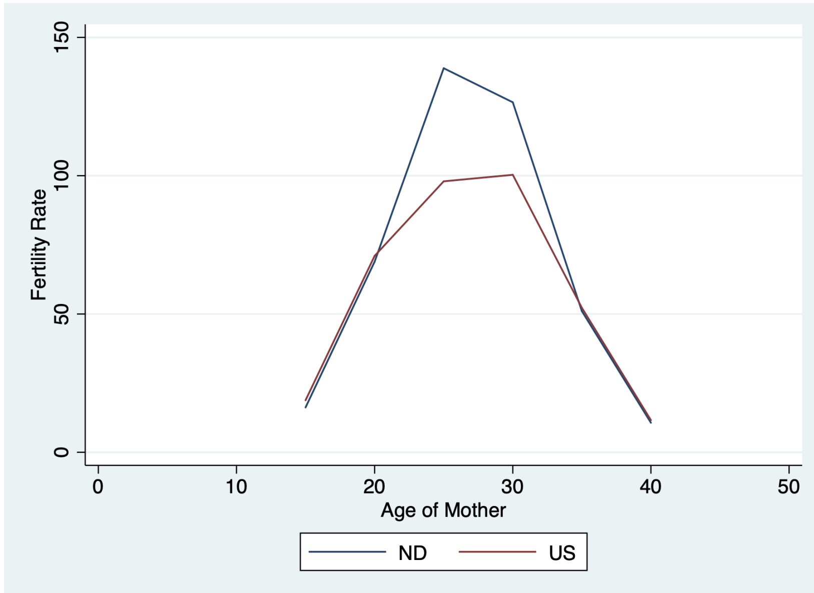


Figure 4: Fertility rate by age of mother, US & ND, 2017

## 2 Economic Change & Social Change

There is no shortage of research on the connections between economic changes and behavioral changes. However much of the literature seems preoccupied with negative shocks to wealth, employment, or income, and associated changes in behaviors such as crime or fertility, such as Sobotka et al. (2011). While important from a policy perspective it would seem to be equally important to understand behavioral responses to positive shocks, especially since the response to positive shocks would not necessarily be symmetric with response to negative shocks. Investigations of this type of result seem to be underrepresented in the literature.

Alam and Portner (2018) looked at income shocks in Tanzania and the timing of fertility. The timing issue is very important because the availability and timing of financial resources, while not changing the total cost of raising a child, can impact the ability to afford a child, or children. This alone could impact both the total number of children born to a woman and the timing of when she has the children. Directly related to my study, Nassirpour (1985) looked specifically at oil revenue in Iran and the impact on fertility in an effort to explain the persistence of higher fertility despite general "sociodemographic theory" predictions that the fertility rate should decline. A major finding was that increases in revenue to provinces did not contribute to the higher fertility rate.

The academic literature also addresses changes in North Dakota as a result of oil development. Recently James and Smith (2017) looked at crime rates in U.S. counties to determine the impact of shale development while attempting to control for changes in labor market dynamics and migration behavior. They found that "[w]estern North Dkaota was uniquely impacted by the development of shale oil in the mid 2000s" and the shale oil effects "much higher than those for the national sample...". Shale oil has long been a topic within the demography literature and can at least be traced backed to Moen (1984) with the discussion of difficulties forecasting accurately the population in the presence of shale devleopment.. In the case of North Dakota, there is a deviation from the national trend in terms of income, making this a contrast to the Alam and Portner (2018).

The broader socioeconomic changes in North Dakota are looked at as well. Rao (2018) chronicles events in the Bakken region through her encounters with different people attracted into the region for the financial windfalls and fresh starts offered. The diversity of individuals in terms of background makes clear that many aspects of work and life in North Dakota changed as a result of the shale oil boom, and continue to change. The nature of social services practice changed as well as access to medical care (for example see Heitkamp and Mayzer (2018) or Jayasundara et al. (2016)). It is beyond the scope of the current paper, but the rise of the opioid epidemic at this same time created additional problems for both the social work industry and medical services while they were both attempting to deal with the changes from oil.

# 3 Data

Data availability and quality are constant issues in this type of work, especially at the county level. The mixing of data sets from federal agencies like the Energy Information Administration, and state agencies such as the North Dakota Department of Mineral Resources and the North Dakota Department of Health creates numerous challenges. Aside from the issues of data accuracy there are always concerns about the extent to which the data actually match up.

The North Dakota Division of Vital Records (2017) provides data on births by county for resident mothers annually going back to the 1930s. Using decennial census information the department provides crude birth rate information as well. I pair the births data with Census Bureau county intercensal population estimates to derive both an alternative crude birth rate as well as a fertility rate.[2] The raw population data provides the base for the crude birth rate while I use data on women in child bearing years, aged 15 to 45, to calculate the fertility rate.

There are some concerns with population data. In the situation of the Bakken oil play population was difficult to track accurately.[3] US Census Bureau estimates will always lag the actual data but it became clear that there was a significant shortfall if you compared data such as labor force to population. By the time estimates published they were grossly inaccurate. The stock measurements were inaccurate, but so were the flow measurements. Some of this is more based on definition than count. There were many oil field workers that worked in the Bakken fields but never fully transitioned to being a North Dakota resident. Many had spouses and families residing in other states that never migrated to North Dakota.[4] Some of the challenges in this analysis are mentioned in the data resources (see for example Division of Vital Records (2017) wherein it discusses the need for adjustments to populations at risk, though there it refers to mortality data).

The availability of county level control variables over a long enough period of time is, so

[2]For example the 1970s data come from Census (2004).

[3]An oil "play" refers to a collection of oil fields in the same region with the same set of geological circumstances.

[4]It is the case that this means the count of births related to the Bakken oil boom may in fact be undercounted if spouses stayed in other states while workers were in North Dakota. Children born in these other states will not be counted in this data. It seems unlikely this problem can be resolved with existing data sets.

far, a limiting factor in the data. I use the county level personal income measurement by the Bureau of Economic Analysis as a control variable. It is available from 1969 to 2017 and I use it in percentage change terms. The oil price data is the first purchase price for North Dakota oil reported by the Energy Information Administration.

In the next two sections I discuss how the data fit into the various econometric specifications used, for example levels versus percentage changes. There is little clarity about whether people look at changes or levels in terms of their decision process about fertility plans, and any changes to those plans.

## 4 State Level Effects

Figure 1 illustrated the significant increase in oil production for the state. It seems a natural place to start to examine whether there are signs, in a statistical sense, that the increase in oil activity impacted the fertility rate for the state as a whole. The independent variable for evaluation at the state level is the fertility rate, calculated as discussed above. I run several models with modifications to the fertility rate. Using the augmented Dickey-Fuller test we fail to reject the null hypothesis of the presence of a unit root. The first difference of fertility strongly rejects the null of a unit root present. Despite the possibility of a unit root I do run some regressions for state fertility rate in the levels. The covariates for the model include state annual oil production in millions of barrels, the first purchase price for North Dakota oil available from the Energy Information Administration, and the percentage change in personal income from the Bureau of Economic Analysis.

The general model estimated is

$$fertility\ rate_t = \beta_0 + \beta_1 oil\ output_t + \beta_2 oil\ price_t + \beta_3 personal\ income_t + \epsilon_t \tag{1}$$

There are some concerns about high correlations between the oil price and quantity variables, which we would naturally expect. As a first step I regressed oil output on the price variable and captured the residuals. I substituted these residuals in Eq. (1) for the oil price variable to eliminate correlations between the oil price and output variables in the regression. In the estimation I use the level for the fertility rate, the natural logarithm of the fertility rate and the first difference of the fertility rate. The results of the regressions are included in Table 1.

Table 1: Regression Results for General Model

| | *Model S1*<br>**fertility (level)** | *Model S2*<br>**fertility (log)** | *Model S3*<br>**fertility (difference)** |
|---|---|---|---|
| oil output | .0302 | .0005 | .0171*** |
| oil price | .0026 | -.000015 | -.0188** |
| personal income | .2678 | .0037 | .0070 |
| constant | 65.0501*** | 4.1669*** | -1.7149*** |
| $R^2$ | .1957 | .2005 | .3010 |
| $AIC$ | 276.666 | -54.268 | 158.937 |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

There are questions about the proper specification for the impact of the oil variables on fertility. Does the level of oil output impact the level of fertility? Or is it better thought that the level of output will impact the change in the fertility rate? The same questions arise when considering the price and income variables. From Table 1 there is clearly some significance (and clearly more work to do). In Model S3 there is a positive correlation with oil output implying that higher levels of oil output would lead to significant changes in the same direction for the change in the fertility rate. For prices there is a negative sign on the coefficient indicating that higher prices will reduce the change in the fertility rate. The mixing of changes in the dependent variable with levels for the independent variables would require an increase of oil output above the mean by 100 million barrels to adjust the change in the fertility rate up by one for the state of North Dakota. I estimated the regression of oil output on oil price this time with both variables in differences and ran the fertility regression again with fertility differenced and oil output differenced. Those results are in Table 2. In the case of Model S4 we see a situation where the quantity of oil is no longer significant though the price level now is. Perhaps more interesting is that the sign switched in this case to positive, implying that as price increases the change in the fertility rate will increase as well. As a robustness check I ran the same models for the crude birth rate in North Dakota and the significance of the results remained principally the same.[5]

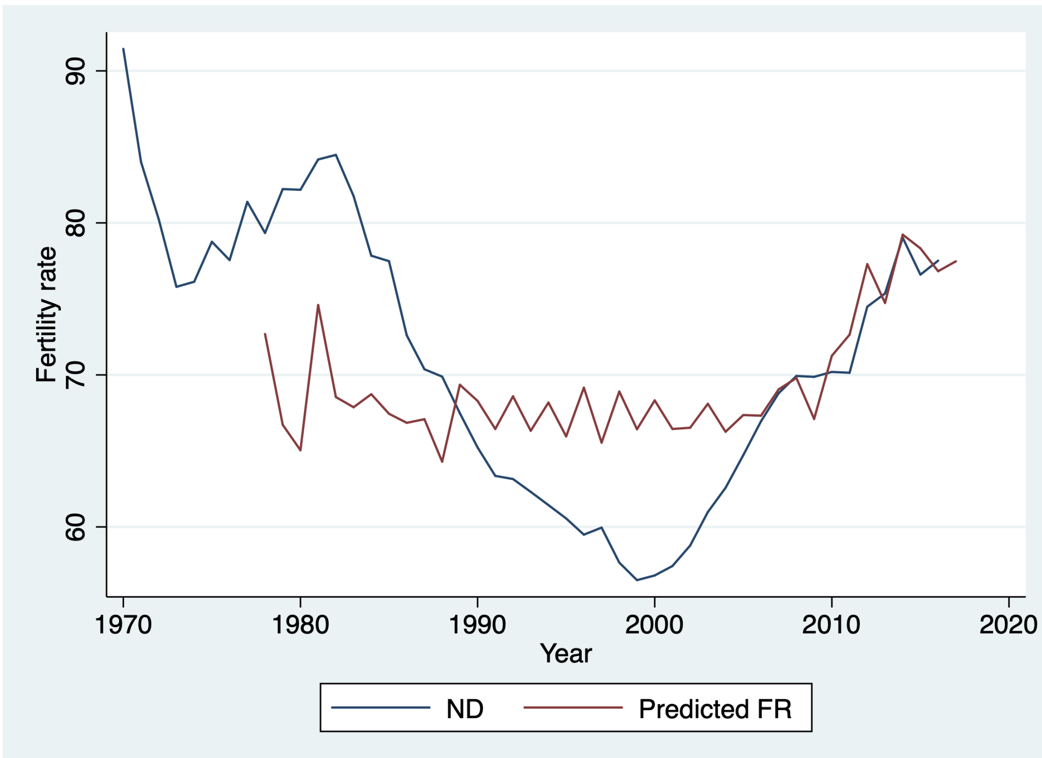


Figure 5: North Dakota Actual and Predicted Fertility Rates by Year, Model S1

From Model S1 I created a predicted value for the North Dakota fertility rate, shown with the actual fertility rate in Figure 6. The model is a naive specification and has some potential statistical issues the figure shows something interesting. From the late 1970s to the mid 2000s the predicted fertility rate did not move much and did not track with the fluctuations in the actual fertility rate. After around 2005 the predicted fertility rate started to perform much better compared with the actual level. This was true of the outcome for state model 4 as well. That the model predicts better as oil increases in importance over

[5]These results are available upon request from the author.

time is not necessarily surprising giving the extent of the oil boom in North Dakota. It would seem then, at the state level, oil price and quantity, depending on the model specification, do have some predictive power for fertility rates.

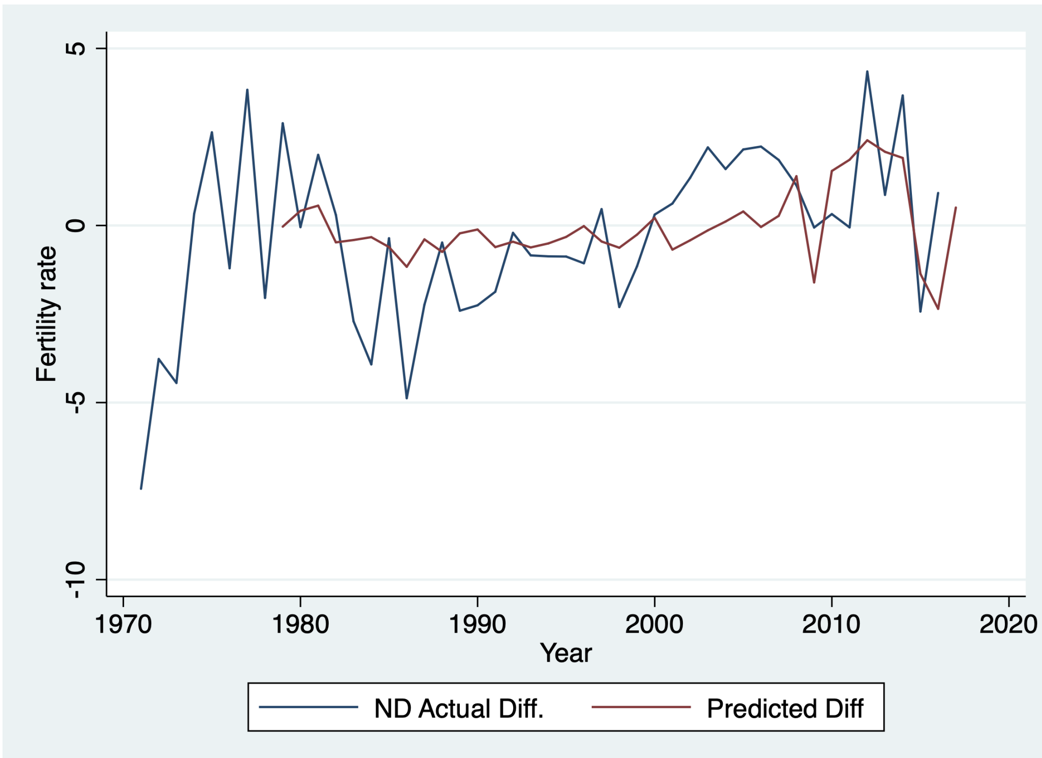


Figure 6: North Dakota Actual and Predicted Differenced Fertility Rates by Year, Model S4

Table 2: Regression Results with All Variables Differenced

| | *Model S4* **fertility (difference)** |
|---|---|
| oil output | -.4688 |
| oil price | .5003* |
| personal income | .0127 |
| constant | 4.3442 |
| $R^2$ | .2477 |
| *AIC* | 157.89 |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

# 5 County Level Effects

At the state level there is some evidence that variables for oil, both price and quantity, are important in determining the fertility rate. Now let's look at the importance of oil in fertility models at the county level for North Dakota. The change of scale to the county level is important. The significant increase in state oil production was not widespread. It was concentrated to a defined region in the West. The *Bakken* formation, as it is called, is found under a group of counties in this region and is responsible for an enormous share of the production in this region, seen from the map in Figure 7. In fact, the publicly available data lists only nineteen of North Dakota's fifty-four counties as producing any amounts of oil from 1951 through 2018. I identify the top oil producing counties in North Dakota and rank them by annual production as a share of total reported North Dakota oil production since the earliest reported data. Dunn, McKenzie, Mountrail, and Williams counties are often identified as the *core* Bakken counties.

Given the geographic concentration, the data on fertility potentially represent a natural experiment with some built in controls given the relative homogeneity of the North Dakota population, especially prior to the discovery of oil. Figure 8 displays the spread of fertility rates across the years for all counties as well as the state as a whole. The dispersion is

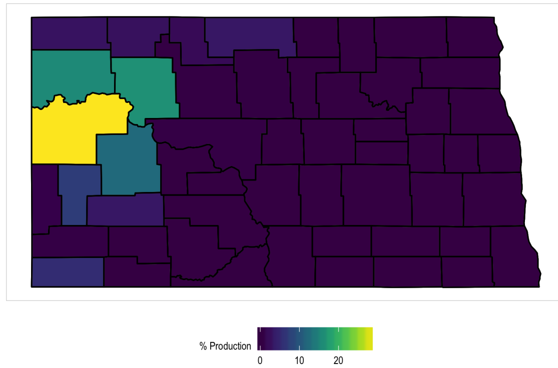


Figure 7: Percent of historical oil production by county, 1951-2018

quite large though the county data roughly follows the pattern of the North Dakota line. In addition I look at the fertility rate of North Dakota as a whole compared to Williams County, North Dakota. Williston, North Dakota, the city at the heart of many stories about man–camps and the ups and downs of the oil boom and bust, is in Williams County. Figure 9 shows that in each of the increases in oil activity, Williams County experienced an increase in fertility rates, significant in both periods. The rate in the most recent oil boom increased steadily, but at a pace much faster than the rate for the state as a whole.

I use a random effects panel specification across counties in North Dakota from the 1970s until 2016 with robust standard errors. The Hausman test failed to reject the null hypothesis that the difference in the coefficients is not systematic. I run the model with two different dependent variables, one with the fertility rate as a level and the other with the natural logarithm of the fertility rate. There were similar concerns with correlation between oil and the North Dakota oil price so once again a I regressed oil output on oil price and captured the residuals for use in this analysis.

The Wald test for all three model specifications strongly rejects the null hypothesis that all model coefficients equal zero. Turning to specific independent variables, the interpreta-

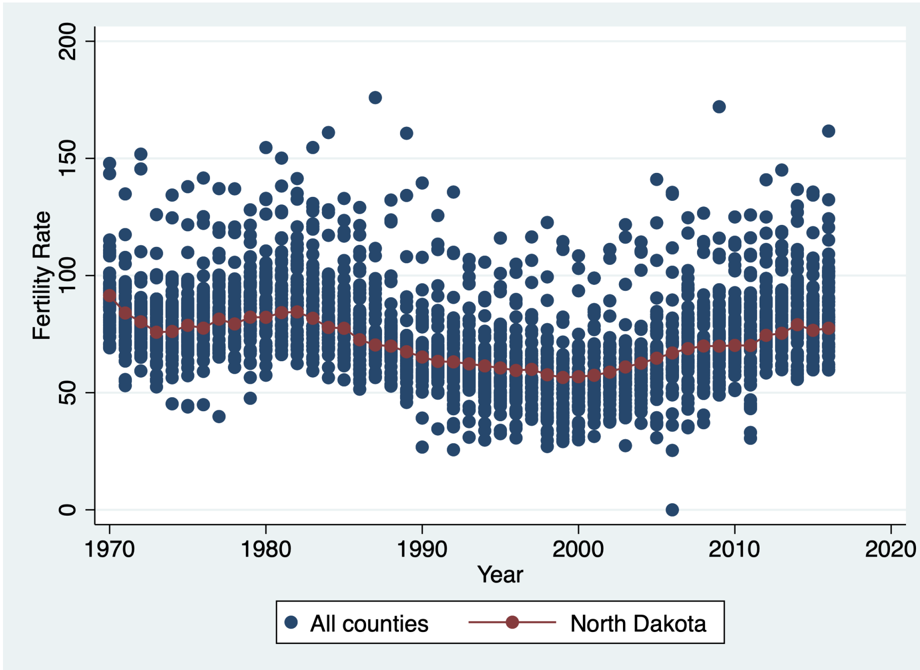


Figure 8: North Dakota State and County Fertility Rates by Year, 1970-2016

tion of the coefficient on oil output is interesting. The average effect of a one percent increase in average county oil output is a small decline in the fertility rate according to all three specifications. The price increase changes sign between Model C2 and C3 with both coefficients being small. The change in sign seems to occur because the fertility rate is in logs and the oil output variable is in either percentage change terms or levels. Further investigation of this is warranted. As in the state level regressions, the percentage change of personal income does not significantly impact the fertility rate.

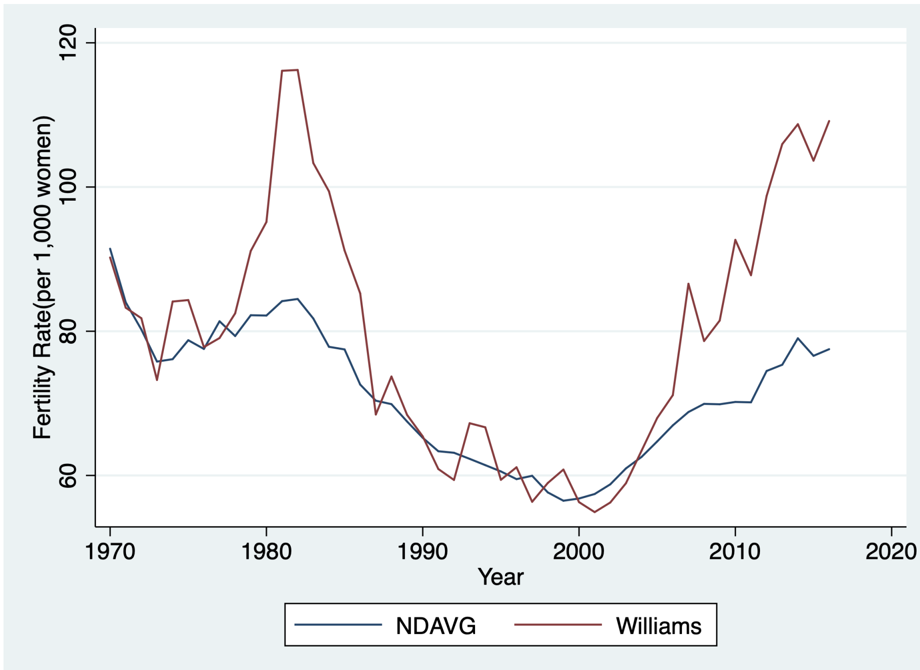


Figure 9: North Dakota State and Williams County Fertility Rates by Year, 1970-2016

Table 3: Random Effects Panel Specification Results

| | *Model C1* **fertility (level)** | *Model C2* **fertility (log)** | *Model C3* **fertility (log)** |
|---|---|---|---|
| county oil output (pc) | .-.0007*** | .-.00001 | |
| oil price | ..2399*** | .0029*** | -.0390*** |
| personal income (pc) | .0218 | .0002 | .00005 |
| county oil (lev) | | | .0419*** |
| constant | 72.8350*** | 4.2512*** | 4.1866*** |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

# 6 Conclusions and Extensions

After decades of declining population in North Dakota there was a reversal that opened up many more business and policy questions than it answered. Unfortunately, there are still many questions left unanswered after many years. For example questions regarding *tempo* and *quantum* in North Dakota and how they might change as a result of shale or national trends (in line with the work of Bhrolchain (2011)) are still elusive. There is some evidence that oil matters for North Dakota state level fertility, and some that it matters for counties too. However the analysis needs more controls to strengthen the results. The labor market and population dynamics for counties and states need to be incorporated to understand better the various factors faced by those making fertility decisions. There is a significant literature on non-marital fertility which may need consideration. The declines in non-marital fertility results found in Schneider and Gemmill (2016) and the economic consequences found in Tach and Eads (2015), may have particular relevance in North Dakota.

Shale oil promises to be a long–term oil play for places like North Dakota and Texas given the results of Smith (2018). The length of time to extract oil will impact the permanency of demographic and social changes as well as compound their effects on populations in the surrounding areas. The economic changes promise to be large as well and the interplay between the social, demographic, and economic changes will be important for policymakers to internalize into their decision processes. For example, spending on health and education may need to adjust to take into account the increased numbers of births. This could include midwives or more hospital capacity and more pre–Kindergarten programs.

Part of the continuing problem is that it is still unclear how individuals interpret increases in economic activity from mining. Is it viewed as inherently temporary or transitory which would bring in results such as Narayan et al. (2008). In addition, the geographic concentration of oil activity, but the diffusion of economic benefits beyond the oil producing counties, sets up a potentially interesting analysis across different types of counties based on income changes. In addition, more, or better, population controls would make the precise rate changes clearer. Along with population change and changes in economic activity there was a shortage of labor in almost all economic and geographic areas. This created a complicated dynamic with growth and income increasing, but less available workers so that growth may have fallen further away from a potential level. In addition, there may need to be geographic control variables added to account for a willingness to commute to certain jobs as well as the clustering of population that occurred in the state.

# References


Alam, S. A. and C. C. Portner (2018). Income shocks, contraceptive use, and the timing of fertility. *Journal of Development Economics 131*, 96–103.

Bacci, M. L. (2011). Low fertility in historical perspective. *Population and Development Review 38*, 72–82.

Becker, G. S. (1991). *A Treatise on the Family*. Harvard University Press.

Bhrolchain, M. N. (2011). Tempo and the tfr. *Demography 48*, 841–861.

Brown, C. (January 31, 2013). North dakota went boom. *New York Times Magazine*.

Census (2004). *Intercensal population estimates by age, sex, and race: 1970-1979*. U.S. Bureau of the Census.

Davey, M. (December 5, 2008). A placid north dakota asks, recession? what recession? *New York Times*.

Division of Vital Records (2017). *North Dakota Resident Vital Event Summary Data, 2002-2016*. North Dakota Department of Health.

Eckstein, Z., T. P. Schultz, and K. I. Wolpin (1985). Short-run fluctuations in fertility and mortality in pre-industrial sweden. *European Economic Review 26*(3), 295–317.

Filoso, V. and E. Papagni (2015). Fertility choice and financial development. *European Journal of Political Economy 37*, 160–177.

Hamilton, B. E., J. A. Martin, M. J. Osterman, S. C. Curtin, and T. Matthews (2015). Births: Final for 2014. *National Vital Statistics Reports 64*(12).

Heitkamp, T. and R. Mayzer (2018). Implications for practice: Risks to youth in boomtowns. *Child Welfare 96*(4).

James, A. and B. Smith (2017). There will be blood: Crime rates in shale-rich u.s. counties. *Journal of Environmental Economics and Management 84*, 125–152.

Jayasundara, D., T. Heitkamp, and R. Ruddell (2016). Voices from the front line: Human service workers' perceptions of interpersonal violence in resource-based boom counties. *Internet Journal of Criminology*, 72–93.

Martin, J. A., B. E. Hamilton, M. J. Osterman, A. K. Driscoll, and P. Drake (2018). Births: Final for 2016. *National Vital Statistics Reports 67*(1).

Martin, J. A., B. E. Hamilton, M. J. Osterman, A. K. Driscoll, and T. Matthews (2016). Births: Final for 2015. *National Vital Statistics Reports 66*(1).

Moen, E. (1984). Voodoo forecasting: Technical, political and ethical issues regarding the projection of local population growth. *Population Research and Policy Review 3*, 1–25.

Narayan, P. K., S. Narayan, and R. Smyth (2008). Are oil shocks permanent or temporary? panel data evidence from crude oil and ngl production in 60 countries. *Energy Economics 30*, 919–936.

Nassirpour, M. (1985). The effect of oil revenue on the fertility pattern in iran, 1952-1976. *Journal of Marriage and Family 47*(3), 785–796.

Oil and Gas Division (2019). *Monthly oil production totals by county.* North Dakota Dempartment of Mineral Resources.

Rao, M. (2018). *Great American Outpost: Dreamers, mavericks, and the making of an oil frontier*. Hachette Book Group.

Roy, T., V. Jayachandran, and S. K. Banerjee (1999, October). Economic condition and fertility: Is there a relationship? *Economic and Political Weekly 34*(42/43), 3041–3046.

Sato, Y. (2007). Economic geography, fertility and migration. *Journal of Urban Economics 61*, 372–387.

Schneider, D. and A. Gemmill (2016). The surprising decline in the non-marital fertility rate in the united states. *Population and Development Review 42*(4), 627–649.

Shahbaz, M., N. Khraief, M. K. Mahalik, and K. U. Zaman (2014). Are fluctuations in natural gas consumption per capita transitory? evidence from time series and panel unit root tests. *Energy 78*, 183–195.

Smith, J. L. (2018). Estimating the future supply of shale oil: A bakken case study. *Energy Economics 69*, 395–403.

Sobotka, T., V. Skirbekk, and D. Philipov (2011). Economic recession and fertility in the developed world. *Population and Development Review 37*(2), 267–306.

Sulzberger, A. (November 25, 2011). Oil rigs bring camps of men to the prairie. *New York Times*.

Tach, L. M. and A. Eads (2015). Trends in the economic consequences of marital and cohabitation dissolution in the united states. *Demography 42*, 401–432.

Tavernise, S. (September 6, 2013). Fertility rate stabilizes as the economy grows. *New York Times*.

Wigniolle, B. (2013). Fertility in the absence of self-control. *Mathematical Social Sciences 66*, 71–86.